

2006

Application of binaural recording in the video game industry

Max Leong
Edith Cowan University

Follow this and additional works at: https://ro.ecu.edu.au/theses_hons



Part of the [Game Design Commons](#), and the [Other Music Commons](#)

Recommended Citation

Leong, M. (2006). *Application of binaural recording in the video game industry*. https://ro.ecu.edu.au/theses_hons/1412

This Thesis is posted at Research Online.
https://ro.ecu.edu.au/theses_hons/1412

Edith Cowan University

Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study.

The University does not authorize you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following:

- Copyright owners are entitled to take legal action against persons who infringe their copyright.
- A reproduction of material that is protected by copyright may be a copyright infringement. Where the reproduction of such material is done without attribution of authorship, with false attribution of authorship or the authorship is treated in a derogatory manner, this may be a breach of the author's moral rights contained in Part IX of the Copyright Act 1968 (Cth).
- Courts have the power to impose a wide range of civil and criminal sanctions for infringement of copyright, infringement of moral rights and other offences under the Copyright Act 1968 (Cth). Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

Application of Binaural Recording in the Video Game Industry

Max Leong

western
australian
academy of
performing
arts



Music Technology, Department of Music
Western Australian Academy of Performing Arts
Edith Cowan University
Perth, Western Australia, Australia
November 2006

A dissertation submitted to Edith Cowan University in partial fulfillment of
the requirements for the degree of Bachelor of Performing Arts (Music) with
Honours

USE OF THESIS

The Use of Thesis statement is not included in this version of the thesis.

Abstract

The video game industry is one of entertainment technologies' largest growing industries. There are thousands of games in the market covering a variety of different genres. There is also a high level technical sophistication in current games along with the game console that runs it. Games are realistic with astonishingly life like graphics. The backing audio and music tracks are performed by leading orchestras and written by renowned composers.

Great technological strides have been made in the areas of graphic and audio design. The end user expects the very best and the video game industry is constantly improving software and hardware technology.

The dissertation will discuss the research, experimentation and ongoing development of the audio technology in the video game industry, specifically the opportunities offered by binaural recording and spacialisation of video game soundtracks and effects. Different adaptations, applications and integration of audio technologies in the gaming arena will be studied with a focus on recording and playback techniques.

The research into binaural applications aims to provide the game user with life-like audio and levels of realism surpassing the current technology commercially available. It explores ways of delivering spacially accurate audio to the gamer with headphones. This reduces the need for multiple speakers and also integrates with the user's current hardware. It also aims to provide new tools for music composers and sound designers to express their ideas. The end user will be exposed to the most spacially accurate audio through headphone playback.

Declaration

I certify that this thesis does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any situation of higher education; and that to the best of my knowledge and belief it does not contain any material previously published or written by a person except where due reference is made in the text.

Signature.....

Date.....

Acknowledgement

I must take the time to thank a number of core individuals whose support and help has been invaluable during my research work. First and foremost, I would like to thank my family. To dad, Ken, Apple and most of all to my mum. I am deeply indebted for all your support, understanding and patience throughout my education. To Emma, I would have never made it without you. I am so lucky to have you in my life. A very big thank you to my good friend Adam Willett who has been the best classmate anyone can have.

I would like to acknowledge all the talented individuals at the Western Australian Academy of Performing Arts whose insights have shaped me into the person I am now. My Composition supervisor Cat Hope, Software and audio with Sam Cutri, my theory and harmony lecturer Donna Brooks, Robert Sazdov, Mike Eastman, Stewart Smith, Jonathon Paget and Anthony Maydwell. All of you along with the great team at WAAPA have given me such a good foundation for me to pursue this research.

To Chris Murdoch and Andy King. Thank you for being so helpful throughout the past 4 years. Chris, your knowledge and skills on sound engineering and recording techniques have been invaluable. You have vast experience in the area and it has been such a great pleasure to learn so much from you.

A final thank you goes to my special instrumentalist Tashana, Duncan and Sacha.

Contents

1	Introduction	1
2	Background	2
2.1	Video Game Industry Overview – Australia.....	2
2.2	Video Game Industry Overview – U.S.A.....	4
2.3	The Gamer.....	5
3	History of Audio in Video Games	6
3.1	1970s.....	6
3.2	1980s.....	6
3.3	Late 1980s and early 1990s.....	7
3.4	Mid to late 1990s.....	7
3.5	2006.....	7
3.6	Issues with Current Audio.....	8
4	Human Sound Localisation	9
4.1	Localisation.....	9
4.2	Interaural Intensity Difference.....	10
4.3	Interaural Arrive-Time Difference.....	11
4.4	The Pinna.....	12

4.5	Direct Sound.....	13
4.6	Head-Related Transfer Functions (HRTF).....	13
5	Binaural Recording	14
5.1	The Dummy Head.....	15
5.2	In-Ear Binaural Microphone.....	15
5.3	Binaural Demonstration.....	16
6	Aims, Design and Method	18
6.1	Design.....	18
6.2	Method.....	19
7	The Recording	20
7.1	Spatial Point of View.....	21
7.2	X-Y Stereo Miking Technique.....	22
7.3	Spaced Pair Miking Technique.....	23
7.4	Binaural Head.....	23
7.5	In-Ear Binaural Recording.....	24
8	Results and Analysis	25
9	Adaptation	28
9.1	In-ear Binaural Adaptation.....	28
9.2	Binaural Head Adaptation.....	30

9.3	Integration.....	30
9.4	Where to go from here.....	31
9.5	Equipment List.....	32
9.6	The Limitations.....	32
10	Summary and Conclusion	33
11	References	35

List of Figures

4.1	Interaural difference.....	10
4.2	The Pinna.....	11
5.1	Dummy Head.....	15
5.2	Detail of Pinna.....	15
5.3	In-ear Microphone.....	16
5.4	Usage of Microphone.....	16
5.5	Binaural Spatial Demonstration (Multi-Media Clip).....	16
7.1	Capture of Solo Performance.....	20
7.2	Setting Camera.....	21
7.3	Recording View.....	23
7.4	Miking Set-up.....	23
8.1	XY Miking (Multi-Media Clip).....	25
8.2	Spaced-Pair Miking (Multi-Media Clip).....	25
8.3	Binaural Recording (Multi-Media Clip).....	25
8.4	Binaural Head (Multi-Media Clip).....	25
9.1	Screenshot of Grand Turismo 4.....	28

Chapter 1

Introduction

The video game world is constantly advancing with technology introducing new and exciting ways of entertaining the gamer. Advances in graphics and audio serve the user by making the game more realistic and playable. The quest for ultra realistic action has seen game console technology improve dramatically. Composers and audio designers can freely express their creative ideas with little boundaries. This is a luxury taken for granted when compared to the monophonic audio of the Pac-man era in the 1980s.

The advancement and availability of portable gaming consoles has increased the use of headphones by gamers. This dissertation researches ways of improving the audio realism for those gamers. Having a sense of space within the artificial game environment (spacialisation) can greatly add to the realness of the game. Multiple speaker configurations have many advantages in bringing a surround space for the gamer, but the cumbersome nature and cost involved can be seen as big disadvantages.

Practical experimentations will be carried to explore ways of minimising the reliance of multiple speaker arrays and obtaining the spacial element in the artificial game environment.

Chapter 2

Background

Video games are part of the lives of countless individuals around the world. Television, music CDs, magazines and billboards all constantly remind us the popularity of video games.

2.1 Video Game Industry Overview – Australia

The Interactive Entertainment Association of Australia (IEAA) released an independent study on the thoughts and habits of Australians in relations to computer and video games.

Here are some of the findings. Of 1601 Australia households surveyed at random in 2005:

1. 76% have a device for playing computer or video games. To divide even further, Out of these 1222 households with the devices;
 - a. 91% have a personal computer
 - b. 57% have a home video game console
 - c. 14% have a hand-held game device
 - d. 42% are homes with children

-
2. Of those who played a computer or video game in the past year, 38% are female.
 3. With regards to the play behaviour;
 - a. 70% Play at least once a week
 - b. Play up to an hour in one go
 - c. Play with others
 - d. Will be playing video or computer games as much or more in 10 years
 4. 40% of game households in Australia purchased a computer or video game in the first half of 2005
 5. 54% Plan to purchase a computer or video game in the second half of 2005
 6. The top selling Games (by unit sold) are;
 - a. Gran Turismo 4
 - b. GTA: San Andreas
 - c. Pokemon Emerald
 - d. Need for Speed: Underground 2
 - e. Simpson's Hit & Run Platinum

(Interactive Entertainment Association of Australia 2006)

2.2 Video Game Industry Overview - United States of America

The Entertainment Software Association (ESA) is the trade association of the computer and video game industry in the United States. It was formed in April 1994 as the Interactive Digital Software Association (IDSA) and was renamed to ESA on July 16, 2003. The following are some interesting statistics based on the ESA analysis of the market.

1. U.S. computer and video game software sales grew four percent in 2005 to \$7 billion - a more than doubling of industry software sales since 1996.
2. The average game *player* is 33 years old and has been playing games for 12 years.
3. The average age of the most frequent game *buyer* is 40 years old. In 2006, 93 percent of computer game buyers and 83 percent of console game buyers were over the age of 18.

(Entertainment Software Association 2006)

Both the American and Australia associations' forecasts show a growing trend in the sales figures as the consumer's demand for new and exciting games increases. The statistics show that video games reach great audiences in our society therefore any advancement in technology will benefit a large number of individuals.

2.3 The Gamer

The Gamer or the end user is one of the most important aspects of the game. This fundamental concept is always on the mind of all game producers, and the reason the game was made in the first place was to service the end user. Without the gamer, there will be no industry.

A realistic game audio enhances the emotion and connection to the games. It is a powerful way of making the game a more enjoyable experience.

The application of binaural recording allows the user to be totally immersed in the surroundings of the game. A new sense of audio realism is explored. When we couple this with the current advances in graphics, the future enjoyment for the gamer is looking very promising.

Chapter 3

History of Audio in Video Games

It is important to research and embellish the history and progression of audio technology in video games. This gives attention to the evolution of the computer game as well as a clue to the trend of where technology is heading. A short overview of the technological advances through the decades is shown below.

3.1 1970s

During the time of its infancy, audio for computer games was all analogue waveforms. They were stored on a physical medium such as compact cassettes. Also in this decade, the first chip music using digital means was implemented. Electrical impulses from a computer code are altered by a specific computer chip where it is then output to the speakers.

3.2 1980s

With the advancement of silicon chip technology along with the fall in the cost, the chips in arcade machines allowed sound generation through different channels. Throughout the 1980s, game consoles were released which were capable of more channels and the ability to simulate instruments for melodies through simple tone generation.

3.3 Late 80s and early 1990s

This decade saw the expansion of early digital synthesis and sampling techniques being implemented by the various console manufacturers. Sound generating chips were still widely used and this decade also brought the 16bit console era to consumers.

3.4 Mid to late 1990s

The technology has moved to pre-recorded and streaming music in the games. The use of optical drives allowed audio to be streamed from the optical disks. Other advancements included 'Sound Card' development for use in Personal Computers.

3.5 2006

Game Consoles are now available with technology such as Digital Dolby, Dolby TrueHD and DTS-HD support, sampling and playback rate of 16-bit @ 48 kHz, hardware codec streaming, and potential of 256 audio simultaneous channels (Computer and Video game music 2006).

The following events draw attention to the popularity of game audio and music;

The first officially sanctioned Final Fantasy concert in the United States was performed by the Los Angeles Philharmonic Orchestra at Walt Disney Concert Hall in Los Angeles, California, on May 10, 2004.

All seats at the concert were sold out in a single day. "Dear Friends: Music from Final Fantasy" followed & was performed at various cities across the United States.

On July 6, 2005, the Los Angeles Philharmonic Orchestra also held a Video Games Live concert, which was founded by two video game music composers at the Hollywood Bowl. This concert featured a variety of video game music, ranging from Pong to Halo 2. It also incorporated real-time video feeds that were in sync with the music, as well as laser and light special effects (Computer and Video game music 2006).

3.6 Issues with Current Audio

The technological advances in computer game audio highlight the progression towards ultra realistic audio. Both visuals and audio aspects of the game need to be constantly improving to show greater levels of realism.

Whilst the current audio has moved leaps and bounds, the spacial aspect is still limited to hardware in which the consumer can afford, resulting in poor playback and unrealistic effects. The need for a standard surround sound array is still very much a reality for many gamers wanting the full surround experience. This set up is not portable for the end user.

This dissertation will look at ways of presenting the audio with accurate spacial features without the use of an array of speaker set-ups. It looks for an economical, efficient technique of presenting spacial audio for the gamer.

Chapter 4

Human Sound Localisation

In order to reproduce the most accurate representation of real-life surround audio, we must first investigate the unique ability of human ear localisation. The human ear can detect and localise sounds from all directions around them. There is much to be gained by exploring the workings of nature and applying it to the field of game audio. By understand the way humans localise sound, we can use technology to emulate it.

4.1 Localisation

Localisation is best described as the ability of being able to pinpoint a sound source within an acoustic space as a result of having two ears.

This effect results from the following three cues received by the ear (Huber and Runstein 2001)

1. Interaural intensity difference
2. Interaural arrival-time difference
3. The effects of the pinnae (Outer ear)

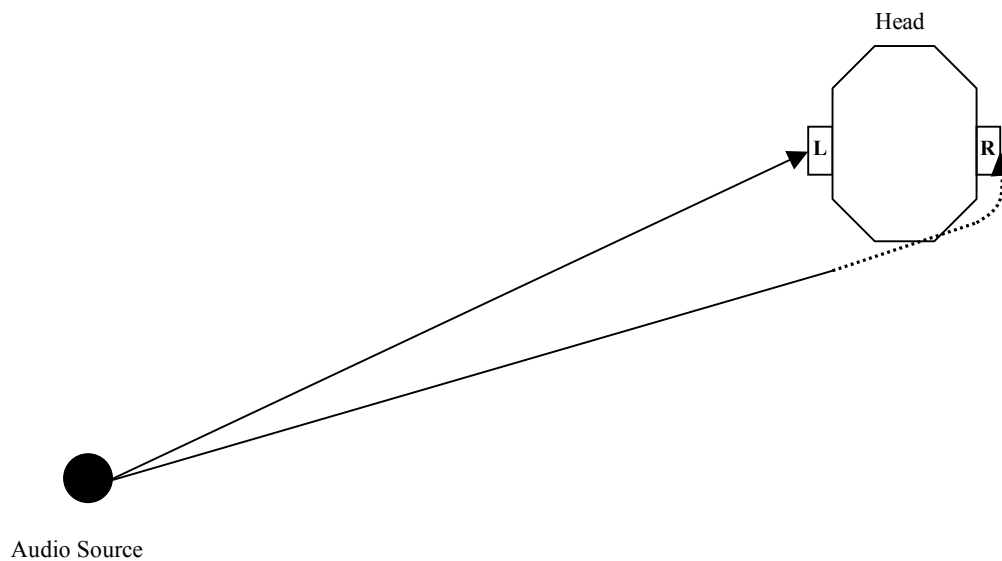


Figure 4.1 Interaural differences

4.2 Interaural Intensity Difference

Interaural intensity difference is an intensity based location method. Middle to higher frequency sounds originated from the left side will reach the left ear at a higher intensity level than the right ear, causing an interaural intensity difference (Figure 4.1). That is, sounds coming from your left hand side will be louder in your left ear compared to your right.

This difference occurs because the head casts an acoustic block or shadow, allowing predominantly reflected sounds from surrounding surfaces to reach the opposite ear.

Interaural intensity difference is relatively insignificant at lower frequencies, where wavelengths are large compared to the head's diameter and the wave easily bends around its acoustic shadow.



Figure 4.2 The Pinna

4.3 Interaural Arrive-Time Difference

Interaural Arrive-Time difference is a different method of localisation employed at lower frequencies. The path length of the sound is slightly shorter to the left ear when compared to the path length to the right ear (Figure 4.1). The sound pressure therefore arrives at the right ear at a later time than the left ear.

Both method of localisation helps to give us lateral localisation cues over the entire frequency spectrum. Both these mechanisms highlight the possibility of confusion between the direction of a sound source at any given angle in front or behind the listener, since both the timing and level differences would produce the same results for both directions. To overcome this ambiguity, an automatic reflex action causes us to instinctively turn or tilt our heads slightly and the resulting changes in timing and level immediately resolve the confusion (Robjohns 1997).

4.4 The Pinna

Both the above-mentioned methods enable us to perceive the angle from which the sound is originating due to intensity and delay cues. However, they do not allow one to determine if the sound originates from the front, behind or below.

The Pinna is the outer visible part of the ear constructed from fibrous cartilage (Figure 4.2). The pinna has two ridges, which produce time delays from the original signal.

Ridge one indicates horizontal sounds. Ridge two indicates vertical sounds.

As sounds arrive at the outer ear, some of the sound enters the ear canal directly, while some is reflected off the curved surfaces of the outer ear and into the ear canal. Since the reflected sound has to travel fractionally further, it is delayed, and in combining with the original sound, produces a comb-filter effect, resulting in characteristic peaks and notches in the frequency response. These frequency-response anomalies depend on the particular direction of sound arrival, and it is thought that we build a 'library' memory of the comb-filter characteristics, which can be used to help provide crude directional cues. (Robjohns 1997)

4.5 Direct Sound

Direct sound takes the shortest path and therefore arrives at the ear first. It provides the ear with information about sound location, size and timbre. The direct sound and reflections within 2 millisecond contribute to localisation. 5-35ms after the direct sound influence image broadening.

4.6 Head-Related Transfer Function (HRTF)

Describes how a given sound wave input is filtered by the diffraction and reflection properties of the head, pinna, and torso, before the sound reaches the eardrum and inner ear. HRTF aids in the neural determination of source location, particularly the determination of source elevation.

HRTFs can also be used in acoustics to create the impression of surround sound without multiple speakers (Head Related Transfer Functions 2006)

Chapter 5

Binaural Recording

Binaural recording produces the most accurate stereo imaging for headphones. It is a recording technique that mimics our natural ability to hear and localise sounds within our space. In recording material the way we naturally hear it as closely as possible, using information on human localisation, we can truly offer the gamer an unmatched level of realism.

Binaural recording works on the following premise:

When we listen to a natural sound source in any direction, the input to our ears is just two one-dimensional signals. That being, the sound pressure at the eardrums. (Bartlett 1999)

Imagine your ears as being a left and a right microphone. Whatever sound you hear around yourself will be recorded. If your ears can tell you which direction a sound is coming from, so can the binaural recording.

Binaural recording with headphone playback is the most spacially accurate method now known (Bartlett 1999).



Figure 5.1 Dummy Head



Figure 5.2 Detail of Pinna

5.1 The Dummy Head

Binaural recording starts with an artificial or dummy head. Figure 5.1 illustrates an example of a dummy head, the KU100 by Neumann. This is a model of a human head with a flush-mounted microphone in each ear. The detail of the pinna is shown in Figure 5.2. The dummy head is built from selected acoustic materials to emulate the sound-transmitting characteristics of a real human head.

5.2 In-Ear Binaural Microphone

The *in-ear* binaural microphone (Figure 5.3) has developed through advances in microphone technology with respect to the quality and size. The *in-ear* binaural microphone is small enough to fit in the ear of the subjects (Figure 5.4). Whatever the subject hears, the *in-ear* binaural microphone will pick up. This is an efficient, portable and economical form of recording.



Figure 5.3 In-ear microphone



Figure 5.4 Usage of microphone

5.3 Binaural Demonstration

The following demonstration (Figure 5.5) is made to highlight the spacial qualities of a binaural recording. The listener should be advised that playback with headphones is recommended.

The demonstration is made with a constant pulse and constant volume intensity. The sound source is at a constant at one meter away and will be circling the listener.

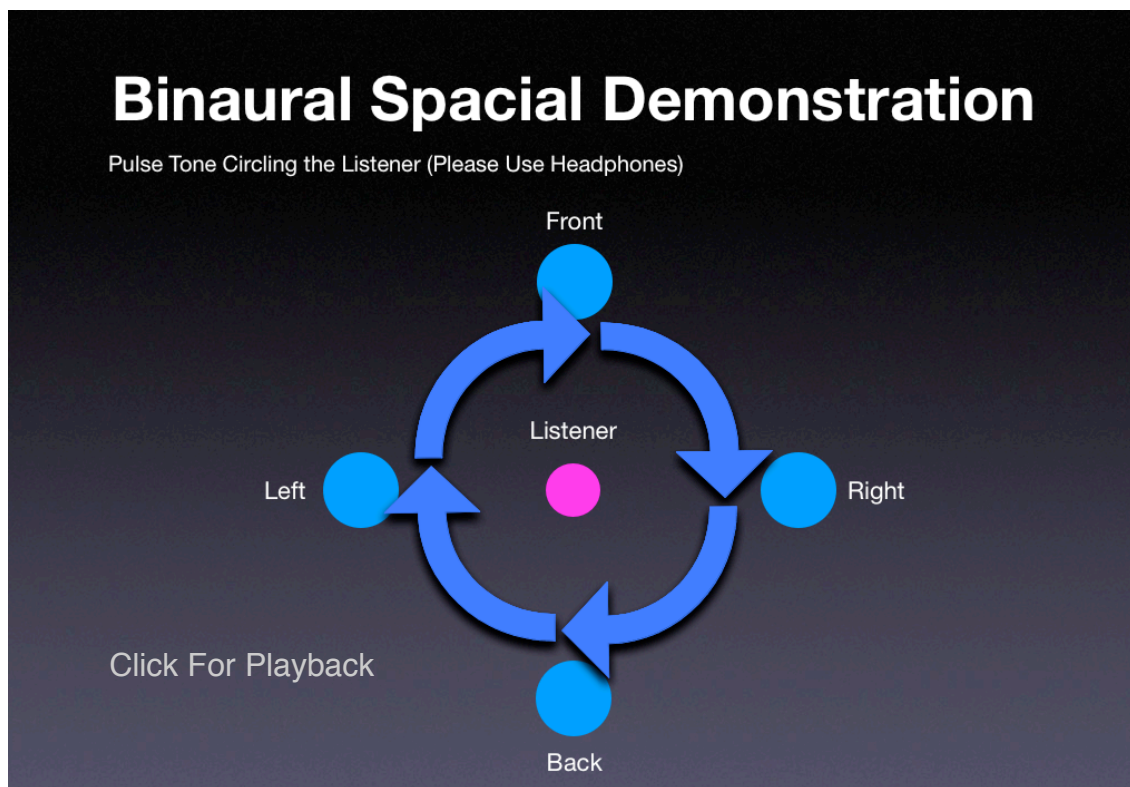


Figure 5.3 Binaural Spatial Demonstration

Chapter 6

Aims, Design and Method

The intended aims of the research dissertation are listed below.

- Development and experimentation of binaural recording technology in the gaming industry
- Offer the user the most accurate way of listening to audio
- Methods to integrate new technologies with the user's current hardware
- Ways to provide the gamer with life-like audio and high levels of realism
- Provide new tools for audio composers to express their ideas

The primary aim however, is to learn and discover new ideas and techniques during the course of the experiments.

6.1 Design

The objective of this chapter is to design a comparative view between conventional and binaural recording techniques. An accurate audio representation of the environment is essential. This adds a further level of realism to the entire gaming experience. This design stage tests the viability of the binaural recording method in use for the audio in games.

6.2 Method

An efficient method for the series of recording experiments must be developed. It should be portable and economical. Most importantly the method development gives the opportunity to build on this research in the future. The method gives details of the procedures involved in the practical aspect of the dissertation. It can be seen as a recipe for the binaural recording process.

Chapter 7

The Recording

A recording atmosphere that allows the listener to hear differences between conventional and binaural recording will be explored. The recording captured the solo performance of 3 (three) classical musicians (Figure 7.1). If the results are favourable on this subtle scale, then the same design and method can be adapted for the audio for video games. The focus and attention can be isolated when only one instrument is used. Any pronounced differences between the conventional and binaural recording will be captured. Several takes from different point of views will be made to highlight the adaptability to video game play (Figure 7.2). The use of different recording techniques will be compared with the binaural recordings.

A recording from the instrumentalist's point of view will be captured using the *in-ear* binaural microphone. This gives a very personal account of what the instrumentalist is hearing when they are performing. The performers hear their instruments and the environment completely differently to what the audience hears. The sound of the finger hitting the fret, the breath before the next passage, the sound of the bow hitting the strings are all intricate and personal sounds that only the performer can hear. The *in-ear* binaural recording gives the most accurate representation from a first person's point of view.



Figure 7.1 Capture of Solo Performance

Spacial Point of View

To present a recording from the audience point of view, a binaural head was placed in front of the performer to give the listener a chance to hear the performance from the best vantage point possible. This method can have the potential of giving the gamer a third person perspective (Figure 7.3).

It is important to compare the binaural recordings with the best current recording technique available (Figure 7.4). The stereo recording techniques used for the tests are called the *XY stereo miking technique* and the *Spaced stereo miking technique*. This will give us a good comparison between the traditional methods and the binaural recording method.



Figure 7.2 Setting Camera

The results gained from this experiment will be useful in adapting the recording technique to game audio.

7.2 X-Y Stereo Miking Technique

The *XY technique* is an intensity-dependent system that uses only the cue of amplitude to discriminate direction. With the XY coincident-pair technique, two directional microphones of the exact same type, manufacture and model are placed with their grills as close together as possible (without touching) and facing at angles to each other (generally between 90° and 135°). The midpoint between the two microphones is faced toward the source and the microphone outputs are equally panned left and right (Huber and Runstein 2001)



Figure 7.3 Recording view

7.3 Spaced Pair Miking Technique

Spaced microphones can be placed in front of an instrument or ensemble (in left right fashion) to obtain an overall stereo image. This technique places the two microphones (of the same type, manufacturer and model) anywhere from only a few feet to more than 30 feet apart (depending on the size of the instrument or ensemble) and uses time and amplitude cues in order to create a stereo image (Huber and Runstein 2001)

7.4 Binaural Head

A binaural head microphone was positioned directly in front of the performer. The audio played back to the listener would be one from the best vantage point in front of the performer. This audio can be considered as the 'best seat in the house'. This is very adaptable to as a third person point of view game play.



Figure 7.4 Miking Set-up

7.5 In-Ear Binaural Recording

The majority of individuals outside the music academy would lack the opportunity to hear what it is like playing an instrument. The in-ear binaural microphone lets the listener hear the performance from the performers point of view. We can adapt the same techniques towards game audio.

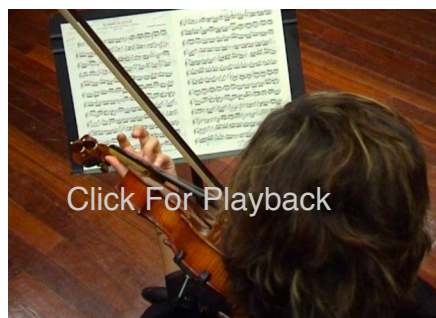
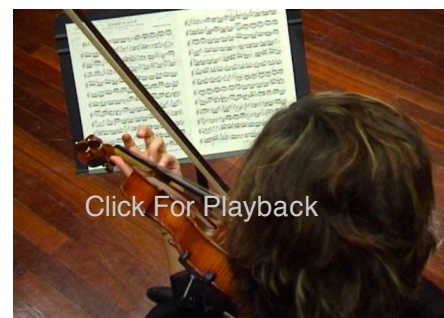
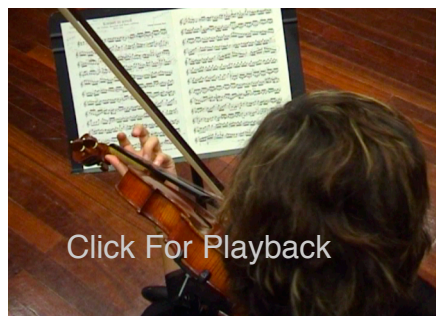
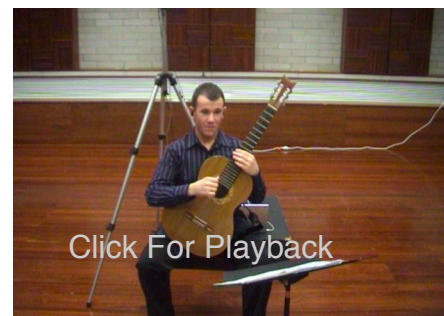
Chapter 8

Results and Analysis

The results and findings were consistent along all three instruments recorded. The violin performance recorded will be used as the demonstration piece for the *in-ear* method whilst the guitar performance will be used for the binaural head method.

Figure 8.1 shows a clip with the audio from the traditional *XY miking recording* technique. Figure 8.2 is a clip from another traditional recording technique, the *spaced pair* microphone technique. These are very familiar recording methods and are common to our ears. They sound similar to other recordings using these techniques and therefore can be used for the basis of our comparison.

Figure 8.3 is a clip using the *in-ear* binaural microphones. The first thing you notice is the different spacialisation aspect compared to the standard recordings. A more accurate recording was made from the violinist point of view. Having the microphones in the ear of the performer gives a very intimate insight into what the instrumentalist is hearing around them. This gives a true first person account from the ears of the performer herself.

**Figure 8.1** XY Miking**Figure 8.2** Spaced-Pair Miking**Figure 8.3** Binaural Recording**Figure 8.5** Binaural Head

Here are some interesting things to take note on Figure 8.3 (Binaural Recording). Some of the very astute listeners might already have picked this up.

- Notice that the left side of the audio is louder. This is because the violin is held on the left hand side and the F-holes are directly below the left ear.
- You can also hear the string hitting the wooden neck by the pressing of the fingers.
- The very astute listeners would be able to make out the intermittent clicking sound on the right hand side was due to a bangle worn on the right wrist.

All these intimate and personal sounds were recorded with good spacial accuracy. This means that if the instrumentalist hears sounds coming louder from the left then this translates to the eventual listener hearing the audio coming louder from the left.

To put it very simply, whatever sounds the performer can localise in the space around them, the eventual listener will be able to localise as well.

The very first ears that the sound enters are the player's ears herself. To put it another way, the first person to ever hear the sound coming out from the violin will be the player herself. We get a very special opportunity to have an audio insight from the first person aspect.

Figure 8.4 show a clip from the third person point of view. The binaural head was placed in front of the performer to capture the audio from the best vantage point in the auditorium. This captures the ambient environment accurately into audio.

Chapter 9

Adaptation

The whole recording process and experiments gave very positive results in spacialisation for the listener. It shows accuracy in determining sounds around the first person's space. This chapter investigates the different possibilities of adapting the research of binaural recording to the video game arena.

9.1 In-ear Binaural Adaptation

The last chapter showed results from the *in-ear* binaural recording method.

This form of recording can be adapted and be used in the first person computer game genre. Lets take a first person car racing game as an example (Figure 9.1). By making a binaural recording of a real racing car driver wearing the microphone, we can have an accurate representation of what the race will sound like from the driver's first person point of view. This is a privilege that only a few can experience.

There will be sounds from below when his foot slams the brakes. There will be sounds from the gear stick to the left as he quickly changes to 2nd gear. There will be sounds from outside when another car overtakes him on the right.



Figure 9.1 Screenshot of Gran Turismo 4

There will also be the constant engine sound. All these intricate and personal first person sounds are captured.

Lets now take that recording from the real racing car driver's point of view and apply it to the computer game. When playing the game, what you hear through your headphones will be exactly what a real racing car driver is hearing during a race. This will give a whole new level of realism to the artificial environment.

Whilst many console games offer some level of spacialisation for the gamer, the application of binaural recordings will offer a far more developed atmosphere and sense of space for the gamer.

9.2 Binaural head Adaptation

The binaural head was placed to record from the best vantage point of the performance. This same technique can be adapted to game audio. The binaural head can be placed in the best vantage point to watch the car race. This can be from the stadium stands, from the roadside or from the pit lanes. The whole atmosphere of the race from the crowd's point of view can be captured. The binaural head technique can give specialisation to the gamer capturing the live ambience of the car race as it progressed.

9.3 Integration

Binaural recording allows us to record audio in the same way the human ear hears and localises sound. We can then reproduce and playback that same recorded audio to simulate the environment it was recorded from. This form of technology and recording works best when played back on headphones.

Binaural recording can be easily integrated and played back in current hardware game systems. The majority of gamers, especially those with a portable hand held console use headphones whilst playing. The binaural method integrates to their existing systems with little to no more expense. The gamer can enjoy the benefits of the new application immediately.

Binaural recording application can give a surround aspect to the game lessening the need for physical speakers to be placed around the listener.

9.4 Where to go from here

Future developments can involve specific live field recordings to mimic the environment of the game. If the game was set in an abandoned warehouse then the field recording will be made in a specific abandoned warehouse. An actor with an *in-ear* binaural microphone will be able to act certain game scenes as the director sees fit. Let's say the game requires the audio of a scene where the hero is being chased by the enemy through a maze of stairs. The actor can act this scenario very easily with the recording device on. The audio will have the ambient noises of the warehouse along with all the accurate sounds that occur during acted chase scenes. The audio can then be integrated to the video game.

A multi-player system can also be implemented. For example, using 10 professional basketball players each with an *in-ear* binaural microphone. The recording will be made on location in a real court with an audience. All aspects of the game including every shot, every step and every rebound will be recorded. Each professional player represents a player within the multi-person game. Different scenarios will be created and the corresponding audio captured. The selected audio can then be used in the virtual basketball game to give a new level of realism.

There are countless applications from this research towards the video game arena. It is about the capturing and playback of the most spacially accurate audio for the end user.

9.5 Equipment List

2 x serial matched DPA 4011

2 x serial matched Neumann Series 180 KM 183 mt

Motu 896HD

Mackie HR824

Sony DSR-PD170P

Sony DCR-TRV80E

Panasonic GS400

Soundman OKM II Klassik/Studio A3

Crown Sass-P MK II

Sennheiser HD265

9.6 The limitations

There are several limitations to the equipment used. Firstly, there was an unfortunate loud noise floor that exists in this particular brand of *in-ear* binaural microphone. To the listener, it will sound like a low hum.

The binaural recording process and adaptation to the gaming audio can be greatly improved with better quality microphones. The binaural microphones used in this dissertation were of high quality but the inherent noise floor was an issue. When adapting this to the game arena, the recording must be accurate and without internal noise from the microphones or the machines running it.

Chapter 10

Summary and Conclusion

The dissertation began with looking into the video game industry to gain an insight into current statistics and video game player patterns. The audio history of video games was also explored to look for trends and emerging patterns. Video games are becoming more sophisticated and realistic. The graphic and audio component is advancing along with new generation game consoles.

The aims of the dissertation were then outlined with major goals being the research and development of binaural recording, the integration with current hardware and offering the gamer the most spacially accurate playback. The goal is also to provide new tools for audio composers to express their ideas.

To achieve these aims, research on human localisation was done. The dissertation explored the nature of human localisation as the result of intensity and delay cues, and the effects of the pinnae. This knowledge helped form the foundation for the research of suitable technology to mimic nature more closely.

The dissertation discussed binaural recording as a technology suitable to meet the aims and goals. This technique gives the most spacially accurate playback when used with headphones. A series of recordings was made to compare traditional recordings and binaural recordings. The design and method was tuned to be adaptable for game play.

The recording showed that a very personal and intimate capture could be gained by using binaural recording. Sounds that only the first person can hear are captured with spacial accuracy and it is reflected on the audio that the listener hears. The environment and any sound that occurs are played back to the gamer exactly as the instrumentalist hears it.

The recording results were discussed with the aid of multimedia clips. Clips from the traditional method of recording were compared to the binaural clips. The focus was on the binaural recording and its spacialisation features. The knowledge gained in the recording process will aid in the development of future methodologies.

Binaural recording and its integration in the video game arena were explored. Games that require headphones are the most adaptable for binaural recording. Portable game consoles fall in that category.

The dissertation shows great promise in the area of game audio. Binaural recordings have the potential of giving the user a greater sense of spacialisation in the artificial environment. Further research and experiments will greatly improve the development of this technique. The application of binaural recordings to video games can only improve the enjoyment of the gamer.

References

Banu, Gunel. "*Analysis of Coincident Microphone Array Recordings for Auralization Applications.*" Ph.D. diss., Queen's University of Belfast, 2004.

Bartlett, Bruce, and Jenny Bartlett. *On Location Recording Techniques*: Focal Press, 1999.

Binaural Recording. (April 16, 2006), In *Wikipedia, The Free Encyclopedia*.

Retrieved April 16, 2006, from

http://en.wikipedia.org/wiki/Binaural_recording.

Blake, Tchad. "*Binaural Excitement.*" (1997 December), Sound on Sound. Retrieved April 16, 2006, from

http://www.soundonsound.com/sos/1997_articles/dec97/tchadblake.html

Brandon, Alexander. *Audio for Games : Planning, Process, and Production*: New Riders Games, 2004.

Breebaart, Jeroen Dirk. "*Modeling Binaural Signal Detection.*" Ph.D. diss., Technische Universiteit Eindhoven, 2001.

Cardiff, Janet. *Janet Cardiff: The Walk Book*: Verlag Der Buchhandlung Walther Konig; Book & CD edition, 2006.

Computer and Video Game Music. (April 16, 2006), In *Wikipedia, The Free*

Encyclopedia. Retrieved April 16, 2006, from

http://en.wikipedia.org/wiki/Computer_and_video_game_music

-
- Dummy Head Recording. (April 16, 2006), In *Wikipedia, The Free Encyclopedia*. Retrieved April 16, 2006, from http://en.wikipedia.org/wiki/Dummy_head_recording
- Entertainment-Software-Association (2006). "*Essential Facts About the Computer and Video Game Industry*." Retrieved August 9, 2006, from <http://www.theesa.com/archives/files/Essential%20Facts%202006.pdf>.
- Fish, Jonathan B. "*Interactive and Adaptive Audio for Home Video Game Consoles*." Ph.D. diss., Simon Fraser University, 2003.
- GameSpot PS2 (2006). "Gran Turismo 4". Retrieved October 8, 2006 from <http://au.gamespot.com/ps2/driving/granturismo4/screenindex.html?page=2>
- Head-Related Transfer Functions (2006). In *Wikipedia, The Free Encyclopedia*. Retrieved April 16, 2006, from <http://en.wikipedia.org/wiki/Hrtf>
- Huber, David Miles, and Robert E. Runstein. *Modern Recording Techniques*. 5th ed: Focal Press, 2001.
- Interactive Entertainment Association of Australia (2006). "*Game Play Australia 2005*" Prepared by Bond University. Retrieved August 9, 2006, from http://www.ieaa.com.au/doc/_3.pdf
- Katz, Bob. *Mastering Audio: The Art and the Science*: Focal Press, 2002.
- Klepko, John. "*5-Channel Microphone Array with Binaural-Head for Multichannel Reproduction*." Ph.D. diss., McGill University, 1999.
- Marks, Aaron. *The Complete Guide to Game Audio: For Composers, Musicians, Sound Designers, and Game Developers*: CMP Books; Bk&CD-Rom edition, 2001.

-
- Marshall, Steve. *"Bilocation: Binaural Recording & 5.1 Surround."* (October 2004). Sound on Sound. Retrieved April 16, 2006, from <http://www.soundonsound.com/sos/Oct04/articles/bilocation.htm>
- McCuskey, Mason. *Beginning Game Audio Programming*. 1st ed: Muska & Lipman/Premier-Trade, 2003.
- Owsinski, Bobby. *The Mixing Engineer's Handbook (Mix Pro Audio Series)*. 1st ed: Artistpro, 1999.
- Pinna. (2006). In *Wikipedia, The Free Encyclopedia*. Retrieved March 14, 2006, from <http://en.wikipedia.org/wiki/Pinna>
- Plunket-Research-Ltd. *"Entertainment and Media Industry Trends."* (2006), Retrieved August 9, from <http://www.plunketresearch.com/Industries/EntertainmentMedia/tabid/199/Default.aspx>
- Qian, Jinyu. *"Virtual Sound Localization Using Head Related Transfer Functions Modified in the Spectral Modulation Frequency Domain."* Ph.D. diss., State University of New York, 2005.
- Reid, Lawrence Gregory. *"Active Binaural Sound Localization Techniques, Experiments and Comparisons."* M.Sc. diss., York University, 1999.
- Robjohns, Hugh. *"Stereo Microphone Techniques Explained"* (February 1997). Sound on Sound. Retrieved April 16, 2006, from http://www.soundonsound.com/sos/1997_articles/feb97/stereomiking.html
- Sanger, George Alistair. *The Fat Man on Game Audio: Tasty Morsels of Sonic Goodness*: New Riders Publishing, 2004.

Stereo Techniques (2006). *"Microphone University"*. Retrieved April 16, 2006
from <http://www.dpamicrophones.com/>

Tomlinson, Holman. *5.1 Surround Sound: Up and Running*: Focal Press, 1999.

White, Paul, and Robjohns, Hugh. *"Frequently Asked Questions: Surround
Sound."* (January 2001.) Sound on Sound. Retrieved February 1, 2006, from
<http://www.soundonsound.com/sos/jan01/articles/surround.htm>

Wilde, Martin. *Audio Programming for Interactive Games*. 1st ed: Focal Press; Bk&CD-
Rom edition, 2004.